Mon. Not. R. Astron. Soc. **294,** L41–L46 (1998)

The Lutz–Kelker bias in trigonometric parallaxes

René D. Oudmaijer,[1] Martin A. T. Groenewegen[2] and Hans Schrijver[3]

[1]*Blackett Laboratory, Imperial College of Science, Technology and Medicine, Prince Consort Road, London SW7 2BZ*
[2]*Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, D-85740 Garching, Germany*
[3]*SRON, Sorbonnelaan 2, NL-3584 CA Utrecht, the Netherlands*

Accepted 1997 November 27. Received 1997 November 4; in original form 1997 July 28

ABSTRACT

The theoretical prediction that trigonometric parallaxes suffer from a statistical effect has become topical again now that the results of the *Hipparcos* satellite have become available. This statistical effect, the so-called Lutz–Kelker bias, causes observed parallaxes to be too large. This has the implication that inferred distances, and hence inferred luminosities are too small. Published analytic calculations of the Lutz–Kelker bias indicate that the inferred luminosity of an object is, on average, 30 per cent too small when the error in the parallax is only 17.5 per cent. Yet, this bias has never been determined empirically. In this paper we investigate whether there is such a bias by comparing ground-based measurements with the best *Hipparcos* parallaxes. We find that there is indeed a large bias with an average and scatter comparable to predictions. We propose a simple method to correct for the LK bias, and apply it successfully to a subsample of our stars. We then analyse the sample of the 26 'best' Cepheids used by Feast & Catchpole to derive the zero-point of the period–luminosity relation. The final result is based on the 20 fundamental mode pulsators and leads to a distance modulus to the Large Magellanic Cloud – based on Cepheid parallaxes – of 18.56 ± 0.08, consistent with previous estimates.

Key words: stars: distances – Cepheids – Magellanic Clouds.

1 INTRODUCTION

Although discussed as early as 1953 (Trumpler & Weaver), Lutz & Kelker (1973) were the first to quantify the bias in the absolute magnitude of a star estimated from its observed trigonometric parallax. The principle of the Lutz–Kelker bias (hereafter LK bias) is relatively easy to understand. A given parallax, π, with a measurement error σ_π yields a distance d with an upper and a lower bound. Stars at a smaller distance and stars located further away can both scatter to the observed distance. Since there are more stars outside than inside the distance range – simply because of the different sampled volumes – more stars from outside the distance range will scatter into the distance range than those inside. This effect causes a systematic bias such that measured parallaxes will on average yield too small distances.

The magnitude of the bias can be calculated analytically. Assuming a uniform distribution of stars, LK found that the mean correction to the derived absolute magnitude increases with increasing relative error, reaching -0.43 mag for a 17.5 per cent error in the parallax. The correction itself is considerable, but since we deal with a statistical process describing the numbers of stars scattering inside and outside the allowed distance range, the bias is represented by a probability distribution. This was investigated by Koen (1992), who calculated the 90 per cent confidence intervals of the correction for the bias. He found for the case of a 17.5 per cent error in the parallax (almost a 6σ detection) the same correction as that found by Lutz & Kelker, but derived that the 90 per cent confidence interval ranges from $+0.33$ to -1.44 mag. This is to be compared with the observational error of 0.4 mag based only on the propagation of the error in the parallax. Such a correction is remarkable indeed, and bears the consequence that parallax measurements should be corrected for this effect, or stringent selection criteria in terms of quality of the data should be taken, before the astrophysical interpretation of the data can be performed.

Indeed, it appears that the only way to take into account the bias before deriving astrophysical parameters from parallaxes is to use the tables by Koen (1992) where the correction is given as function of the relative error in the parallax, (σ_π/π). However, as has been pointed out by Smith (1987c), the LK correction to an individual parallax measurement of a star, or to a sample of stars is valid when no a priori additional information, like proper motions or knowledge on the intrinsic absolute magnitude distribution, is available. In the case of a Gaussian or top-hat intrinsic magnitude distribution, Smith (1987a,b,c; see also Turon Lacarrieu & Crézé 1977) showed that the correction is a function of the true absolute magnitude, M_0, the intrinsic spread in M_0, σ_{M_0}, the observed parallax π, and the error on the observed parallax σ_π, and not a simple function of (σ_π/π). For a magnitude-limited sample (complete in parallax to a certain limiting magnitude), this correction, according to Smith (1987c), approaches zero because the combination of the

Malmquist bias and the LK bias lead to a symmetric error in magnitude. For other samples, the corrections reach asymptotically the LK values, which apply when nothing about M_0 or σ_{M_0} is known.

Considering the implications of the above and the fact that the presence of the bias has actually never been established empirically, we devised a simple empirical test using ground-based and *Hipparcos* data.

2 A STATISTICAL BIAS IN PARALLAXES

We investigate here whether there is any change in the absolute magnitude determined from the measured parallax π of a star as a function of the error in the parallax (σ_π/π). If there would be a trend towards too faint absolute magnitudes with larger (σ_π/π), or even a large spread, then it is likely that an LK bias is present.

The *Hipparcos* catalogue is a good starting point to perform such a test. It is hard however, to find a sample of stars for which we *know* their absolute magnitudes from first principles. For example, selecting stars with the same spectral type will not be sufficient, as such a sample will have a large spread in intrinsic magnitudes, will inevitably suffer from misclassifications and have completeness problems. On the other hand, when stars with extremely good *Hipparcos* parallaxes are considered, one may reasonably assume that their distances are well determined. If one then compares such a sample of stars with their (poorer quality) ground-based parallaxes, it is possible to investigate the LK bias. We therefore selected all stars in the *Hipparcos* Input Catalogue (Turon et al. 1993) with trigonometric parallax measurements $0 < (\sigma_\pi/\pi)_{\text{ground-based}} < 1$. From the remaining 4007 objects, we selected those stars in the *Hipparcos* Catalogue (ESA 1997) with

(i) $0 < (\sigma_\pi/\pi)_{\text{Hip}} < 0.05$, i.e. a 20σ detection or better,
(ii) number of rejected data < 10 per cent (Field H29),
(iii) goodness-of-fit smaller than 3 (Field H30).

The first criterion ensures us that we have a sample for which we may hope to assume that the data do not suffer from significant bias problems (if Koen 1992 is correct, the mean bias is at most 0.025 mag, with 90 per cent confidence limits $\lesssim$0.2 mag) while the latter two criteria ensure that the data are not hampered by observational problems as discussed in the accompanying literature to the *Hipparcos* data base. These three extra selection steps left us with a sample of 2187 stars.

The data are plotted in Fig. 1. The upper panel shows the inferred absolute visual magnitude, derived from the ground-based parallax, M_{par}, plotted against the ground-based parallax. The stars follow a well-defined band in the plot. This is easily understood. If one takes the analogy for one object with a measured *V*-band magnitude, it can only follow a straight line when its derived intrinsic magnitude is plotted as function of parallax. Consequently, all objects should lie between the lines defined by the brightest and faintest *V* magnitude in the sample. One can also say that the faintest *V* magnitude in the sample defines a minimum possible value of the parallax for a given value for M_{par}. The difference between the (measured) absolute magnitude of an object and the limiting (i.e. faintest) *V* magnitude of a sample implies a maximum possible distance, and thus minimum parallax. This can be written as

$$\pi > 1000 \times 10^{[0.2\,(M_{\text{parallax}} - V_{\text{max}} - 5)]} \quad \text{mas} \qquad (1)$$

where V_{max} is the faintest magnitude of the sample. A similar relation can be written for the brightest (minimum) magnitude in the sample. The resulting boundaries are indicated by solid lines in

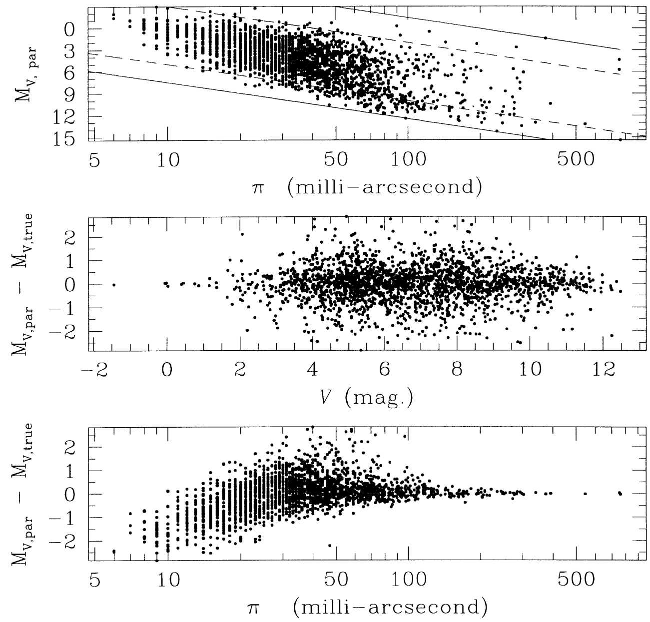

Figure 1. Upper panel: M_{par} as function of π. The ground based parallaxes are listed with integer values. The solid lines give an indication to the zone where we do not expect any data (see text). Middle and lower panel: ΔM as function of V and π. Although there appears no correlation with V, there is a strong correlation with parallax.

the upper panel in Fig. 1, the dashed lines represent $V = 2$ and 10 mag, encompassing the bulk of the sample. The middle panel shows the difference $(M_{\text{par}} - M_0)$, [calculated from $5\log(\pi_{Hipparcos}/\pi_{\text{ground-based}})$, hereafter ΔM] as function of *V*. ΔM shows a large scatter, especially for faint *V*, but the (unweighted) mean $\Delta M = -0.01 \pm 0.8$ mag, which would appear rather reassuring. The propagated errors are not plotted, but these are indicated in Fig. 2.

In the lower panel, ΔM is plotted against π. A strong correlation is present. For large π, ΔM is close to zero – indicating good distance determinations – followed by an increase of the spread in values until ΔM decreases towards *brighter* absolute magnitudes. There are no stars present in the upper left-hand corner of the plot. This is not a real effect, but rather a completeness effect in the data. In reality such objects are further away than expected, and too faint to be included in the sample. This zone is effectively forbidden, as is illustrated by the solid lines in the upper panel, where the lower left-hand corner is void.

Although completeness effects play an important role, the resulting distribution of ΔM is very wide, and illustrates that absolute magnitude estimates based on trigonometric parallaxes are subject to large scatter, with a range of -3 to 3 mag.

Fig. 2 shows ΔM plotted against (σ_π/π). Since the range in σ_π is much smaller than the range in π (the average σ_π is 7.9 ± 2.5 mas), the trends are roughly the same as in the $(\Delta M - \pi)$ relation of Fig. 1. The filled circles indicate the average ΔM, with their standard deviation binned over intervals of 0.1 in (σ_π/π). These data are compared with the results of Koen (1992), in the case of a uniform density of stars and an infinite number of measurements. The thick solid line indicates the mean bias calculated by Koen, and the thick dashed lines represent the 90 per cent confidence intervals of the bias correction. For a smaller number of measurements, the bias and the corrections are larger, while for a decreasing stellar density (Koen's $p = 2$ case), the corrections themselves are somewhat

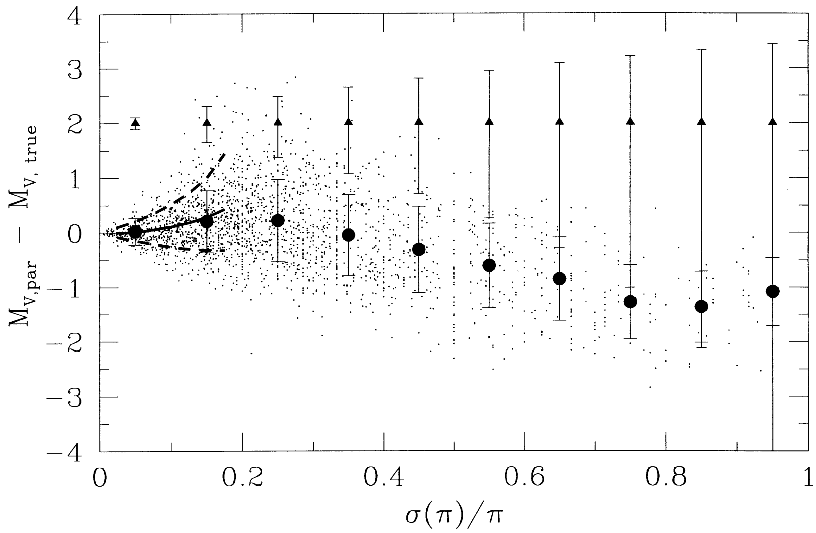

Figure 2. ΔM as a function of (σ_π/π). The triangles with error bars indicate the propagated errors in magnitudes from parallaxes as function of (σ_π/π). The large solid circles indicate the average ΔM binned over intervals of 0.1 in (σ_π/π). The thick lines indicate the LK bias calculated by Koen (1992) for a uniform density of stars. The dashed lines indicate the 90 per cent confidence intervals.

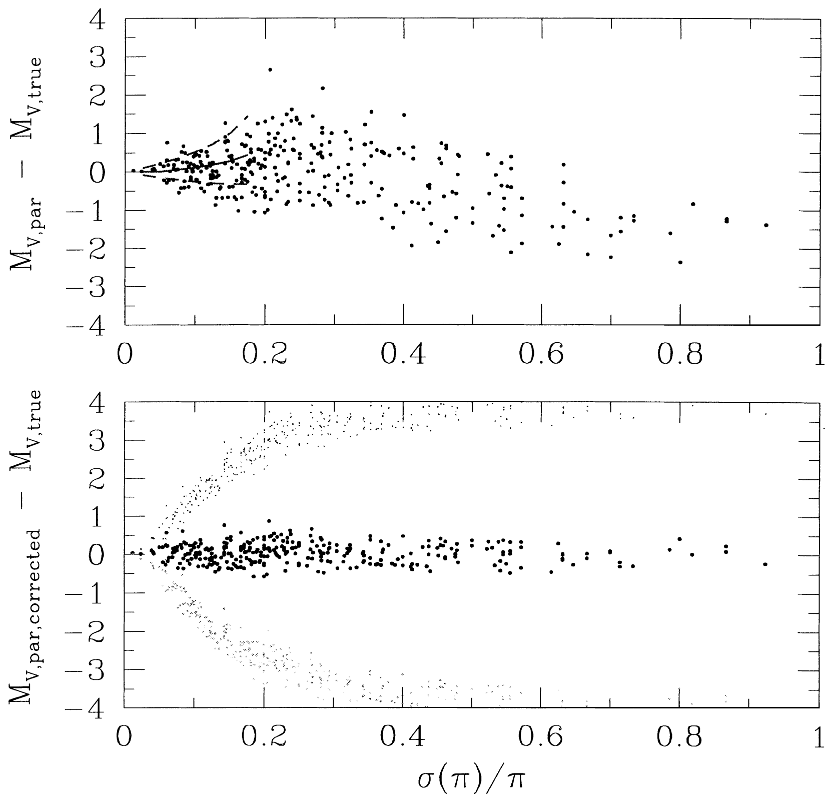

Figure 3. Upper panel: as the previous figure, now for objects with $4 < M_0 < 5$. Lower panel: ΔM with M_{par} corrected using equation (2). The upper and lower clouds of points respectively show the dependence of the solution when a different M_0 is used (M_0= 8.5 and 0.5 respectively instead of 4.49).

smaller, but the confidence intervals are similar. As most of the stars in our sample are located within 100 pc (see Fig. 1), the assumption of a uniform stellar density is probably close to the real situation.

For $(\sigma_\pi/\pi) < 20$ per cent, the observations and the predictions by Koen agree very well, after that, the average ΔM reaches 0, and goes towards too bright magnitudes when (σ_π/π) increases further. It is due to the objects with large (σ_π/π) (which are affected by the completeness of the sample) that the overall mean ΔM is close to zero. If a selection on parallax, or (σ_π/π), had been applied to these data, the resulting mean M_{par} would have led to too faint mean intrinsic magnitudes. Unfortunately, our sample renders the construction of a magnitude-limited sample not possible. It would thus appear that there is indeed an LK-type bias in parallax data. In the following, we will concentrate on the correction for the LK bias.

2.1 A correction for the Lutz–Kelker bias

One of the assumptions in the analysis by Lutz–Kelker (1973) and Koen (1992) is that the absolute magnitude and its spread of a sample of stars are unknown. Smith (1987b,c) and Turon Lacarrieu & Crézé (1977) consider the case of a luminosity function with a Gaussian spread (for a uniform distribution of stars) and derived a formalism for the correction of the LK bias. This correction turns out to be a function of M_0, σ_{M_0}, and (σ_π/π), and converges to the LK value for large σ_{M_0}. A linear approximation, valid when $|\Delta M| <$ 2.17, to this rigorous correction was derived by Smith (1987c):

$$\delta M = \left\{ 1 - \left[\frac{\sigma^2_{M_0}}{\sigma^2_{M_0} + 4.715(\sigma_\pi/\pi]^2} \right] \right\} (M_0 - M_{\rm par}). \qquad (2)$$

For large σ_{M_0}, the linear approximation becomes zero, the exact value of the rigorous correction is close to the LK value (Smith 1987c).

To appreciate the usefulness of this particular result and to see whether the correction can be used as a means to derive the 'true' intrinsic magnitude of a sample of stars, we select a subsample from our sample of objects. To mimic a sample of stars with approximately the same mean intrinsic magnitude, we restrict ourselves to a narrow range of true absolute magnitudes, and apply the correction given in equation (2). There are 313 objects present in the intrinsic magnitude (i.e. derived from the *Hipparcos* parallaxes) bin $4 < M_0 <$ 5. We took the mean M_0, and its standard deviation (4.49 ± 0.28 mag.) as input values for equation (2). The results are plotted in Fig. 3. The upper panel shows ΔM as function of (σ_π/π), which is similar as for the larger sample depicted in Fig. 2. The corrected values are shown in the lower panel, and the mean intrinsic magnitude appears to be retrieved.

Hence, it is possible to correct a sample of objects for the LK bias. It would seem that this correction for the statistical biases is not very useful since one has to know the answer already before applying the correction. However, if one studies a sample of objects of which one may assume that they all have the same M_0, it may be the basis for a powerful method to derive the absolute magnitude. To illustrate this, Fig. 3 also shows the resulting ΔM for other values of M_0 in equation (2). The upper cloud of points was obtained for inserting M_0 = 8.5 in the equation, while the lower cloud of points were calculated using M_0 = 0.5. For $(\sigma_\pi/\pi) \lesssim 0.4$, a strong dependence on the input value of M_0 is present. It appears then, that the correct value of M_0 can be obtained iteratively by varying the input value of M_0 to obtain a horizontal line, or a minimum spread around 0. Following a procedure outlined later in more detail, we found the best value for the intrinsic magnitude of the sample to be in the range 4.48–4.59 when small (< 0.1) and large (> 0.5) values of the (assumed to be unknown) spread in intrinsic magnitude are used, respectively.

The above exercise shows that a correction for the LK bias is not a simple function of (σ_π/π). In the case of an individual object of which nothing is known, the LK correction, along with its large confidence interval is the only remaining option. However, for a sample of which all stars are known to have the same intrinsic magnitude, the LK correction as determined by Smith (1987c) is a potentially powerful tool to determine M_0. As shown above, this method returns a surprisingly good result on our subsample of objects. In the following we will apply this to a sample of Cepheids.

3 THE CEPHEID DISTANCE SCALE

Given the presence of a bias in parallax data described above, it is surprising that several studies based on *Hipparcos* data, whilst not taking the LK bias into account, yield results that are relatively close to previous results. For example, the new zero-point of the Cepheid period–luminosity (PL) relation that Feast & Catchpole (1987, hereafter FC) found, increased the distance modulus to the Large Magellanic Cloud by (only) 0.2 ± 0.1 mag, and a decrease in the, Cepheid-based, Hubble constant of about 10 per cent. Fortunately, we can construct a similar test for the Cepheids as described above. Contrary to the sample in the previous section, where the extremely good *Hipparcos* parallaxes provided the true absolute magnitude, we now have to use another indicator. The true absolute magnitude of the Cepheids is assumed to follow the PL relation for Cepheids (cf. FC)

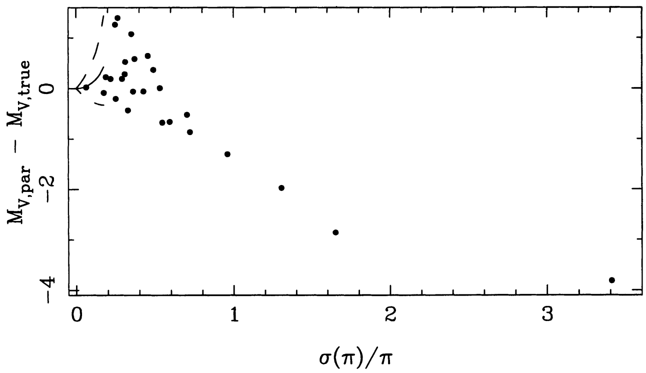

Figure 4. As the previous figure, now for *Hipparcos* measurements of 26 Cepheids. The true magnitude was calculated using the period–luminosity relation (equation 3 with ρ = −1.43).

$$< M_V >= \delta \; \log P + \rho, \qquad (3)$$

with ρ = −1.43 ± 0.10 and δ = −2.81 (as derived and assumed respectively by FC). The *Hipparcos* parallaxes, combined with the reddening corrected $<V>$ magnitudes then give $M_{\rm par}$. We took the data of the 26 'best' Cepheids which contributed the largest weight (87 per cent from a sample of 220 objects) to the solutions from FC. The stars were corrected for reddening in the same way as in FC.

The relation between ΔM and (σ_π/π) is plotted in Fig. 4. The same trend as in Fig. 2 is present. For small (σ_π/π), ΔM tends to indicate fainter intrinsic magnitudes than predicted from the PL relation, whereas for large (σ_π/π), the Cepheids have too bright magnitudes by as much as 4 mag! The fact that the Cepheids with large (σ_π/π) are all too bright is a result of the completeness effect discussed earlier. The reason that the FC zero-point was close to previous values is that the LK bias is cancelled out by the completeness effects to a certain, but ill-defined degree.

One can try to correct for the LK bias using equation (2), provided we know all relevant parameters. A difference with the case illustrated in Fig. 3 is that the parameter we want to estimate is not M_0 but ρ, the difference between M_V and δlog$P$. A second difference is that σ_{M_0} (or equivalently in this case, σ_ρ) is unknown. From the 26 Cepheids we consider the 20 fundamental mode pulsators, as indicated by FC. For this sample FC found a zero-point ρ = −1.49 ± 0.13 (corresponding to a distance modulus to the LMC of 18.76). Three stars have values of ΔM larger than 2.17, and are hence not applicable to equation (2), and will be discarded in the analysis. The reason that we chose not to include the six first overtone pulsators is discussed below. We can now derive an improved value for the zero-point of the Cepheid *PL* relation applying the LK correction to the parallaxes.

Table 1. Zero-point ρ of Cepheids for the 17 best stars.

σ_ρ	ρ	σ (from χ^2+1)	σ (around mean)
0.05	−1.31	2.70	0.0017
0.10	−1.31	0.88	0.0069
0.175	−1.30	0.55	0.020
0.2	−1.29	0.50	0.026
0.3	−1.28	0.40	0.052
0.4	−1.27	0.37	0.083
0.6	−1.25	0.35	0.143

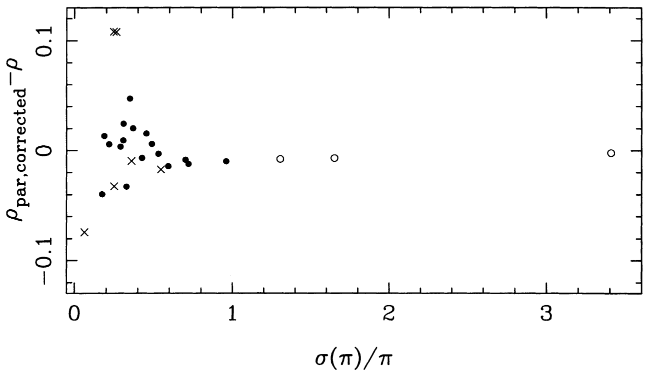

Figure 5. As previous figure, but now for the fundamental mode Cepheids, after correction for the LK bias, for $\rho = -1.29$. The open circles represent the three stars not used in the analysis because they have $\Delta M > 2.17$ (see Fig. 4). The six crosses represent the first overtone pulsators. Note the change in vertical axis compared to the previous figure.

The basic principle behind the following procedure is to iteratively vary ρ until the variance around $\Delta M = 0$ is minimal. We calculated for a range of values of ρ the 'true' M_0 and ρ_{par} from $\rho_{\mathrm{par}} = M_{\mathrm{par}} - \delta \log P$ for every star. The quantity $\Delta\rho = [\rho_{\mathrm{par}}$ (corrected) $-\rho]$ is calculated with ρ_{par} (corrected), the value of ρ_{par} after applying equation (2). Calculated are the mean and standard deviation in $\Delta\rho$ and $Q^2 = [\sum(\Delta\rho/s)^2]/(N-1)$. For s we assumed $\sigma_\rho/\sqrt{N}$, with $N = 17$ the number of stars in the sample. The best value of ρ is found where Q^2 has a minimum, χ^2. The 1σ uncertainty around the best value is estimated from those values of ρ for which $Q^2 = \chi^2 + 1$. The results are presented in Table 1, where also the standard deviation in the mean of $\Delta\rho$ is listed.

A complication is that the spread σ_ρ is unknown. Obviously, if we use $\sigma_\rho = 0$, all corrections will be equal to the observed ΔM. All input values of ρ would yield equally low values of Q^2, and hence the error estimate in ρ based on the variation of χ^2 would be very large. The best value for ρ is almost equal for every adopted σ_ρ. We adopt a final value for $\rho = -1.29 \pm 0.02$, with the error based on the scatter in the best-fitting values of ρ. This means a decrease in ρ of 0.20 with respect to FC for *exactly* the same sample and – all things being equal – an LMC distance modulus of 18.56.

The extent to which this procedure has compensated for the bias is illustrated in Fig. 5, where $\Delta\rho$ is plotted against (σ_π/π) for the case $\rho = -1.29$, $\sigma_\rho = 0.175$ which has a standard deviation around the mean equal to the adopted uncertainty. Note that all 20 fundamental pulsators are plotted, although the fitting was performed excluding the three stars with initially the largest ΔM. The figure also illustrates the reason why we chose not to include the six first-overtone pulsators in the analysis. Three of them give significantly higher residuals. This may be taken as evidence that the procedure by FC to transform first overtone to fundamental mode pulsation period (equations 8 and 9 in FC) introduces additional noise. As FC, we keep the value of δ fixed. Taking the error into account will increase the uncertainty in the zero-point. Repeating the analysis of FC we find for their sample of non-overtone pulsators that changing δ by ±0.06 (the uncertainty in the slope derived by Caldwell & Laney 1991 and adopted by FC) changes their zero-point by 0.05. We performed the same test for our procedure, and find that changing δ by $\pm$ 0.06 results in a shift of the best-fitting ρ of ±0.05. Taking into account errors in ρ ($\pm$ 0.02), in δ ($\pm$ 0.05) and in the visual dereddened magnitudes ($\pm$ 0.06) our best estimate based on a re-analysis of the Cepheid sample of FC is 18.56 $\pm$ 0.08. This value is more in agreement with previous determinations using Cepheids (see e.g. the listing by Madore & Freedman 1997) and RR Lyrae variables (e.g. Alcock et al. 1997).

4 CONCLUDING REMARKS

We have investigated whether there is a Lutz–Kelker type bias present in parallax data. For the derivation of the 'true' intrinsic magnitudes, 20σ parallaxes or better were taken from the *Hipparcos* data, and compared with less accurate ground-based parallaxes that were available for these stars. For small relative errors in the parallaxes, ΔM is distributed asymmetricaly around zero, with a preference for too faint magnitudes, as is expected from an LK-type bias. The spread in values is consistent with the confidence intervals for the bias that were calculated by Koen (1992). For larger values of (σ_π/π), the parallaxes tend to yield too bright magnitudes. This can be explained by completeness effects in the data, where the limiting magnitude of the sample implies a lower limit to the observed parallaxes. A simple method to correct for the bias has been presented and tested. The Cepheid data of Feast & Catchpole (1997) were then investigated, A re-analysis of these data, taking into account any biases, returns a value of the distance modulus of 18.56 $\pm$ 0.08, which is 0.14 mag smaller than FC found, and in good agreement with previous determinations.

Finally, we note that unless parallax measurements are extremely precise, the determination of astrophysical parameters from these data will be affected by LK-type biases. For the moment, either using extremely precise data, or taking into account the LK-bias, with its large confidence intervals, seems to be the only option for individual objects, while the simple correction proposed here, can be used to obtain more reliable estimates of the mean intrinsic magnitude of a sample of stars, provided one knows beforehand that the objects have the same luminosity. The assumption of a uniform distribution of stars used throughout this paper may be improved upon by using Monte-Carlo calculations of the distribution of stars in the line of sight of the targets.

ACKNOWLEDGMENTS

We thank Drs J. E. Drew and L. Lucy for illuminating discussions. Dr H. Smith, Jr is thanked for his constructive comments on an earlier version of the manuscript. RDO is funded through a PDRA grant from PPARC. This paper is based on data from the ESA *Hipparcos* astrometry satellite.

REFERENCES

Alcock C. et al., 1997, ApJ, 482, 89

Caldwell J. A. R., Laney C. D., 1991, in Hayes R., Milne D., eds, Proc. IAU Symp 148, The Magellanic Clouds. Kluwer, Dordrecht p. 249
ESA, 1997, The *Hipparcos* Catalogue, ESA SP-1200
Feast M. W., Catchpole R. M., 1997, MNRAS, 286, L1
Koen C., 1992, MNRAS, 256, 65
Madore B. F., Freedman W. L., 1997, ApJ, 492, 110
Lutz T. E., Kelker D. H., 1973, PASP, 85, 573
Smith H., Jr, 1987a, A&A, 171, 336
Smith H., Jr, 1987b, A&A, 181, 391
Smith H., Jr, 1987c, A&A, 188, 233
Turon Lacarrieu C., Crézé M., 1977, A&A, 56, 273
Turon C. et al., 1993, The *Hipparcos* Input Catalogue. ESA SP-1136
Trumpler R. J. & Weaver H. F., 1953, Statistical Astronomy. Berkeley Univ. Press p. 369

This paper has been typeset from a TEX/LATEX file prepared by the author.